NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER SAN DIEGO, CALIFORNIA 92152

SPECIAL REPORT 75-9 APRIL 1975

LONGITUDINAL STUDIES OF ATTITUDE CHANGE: ISSUES AND METHODS

NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER SAN DIEGO, CALIFORNIA 92152

SPECIAL REPORT 75-9 APRIL 1975

DTIC FILE COPY

LONGITUDINAL STUDIES OF ATTITUDE CHANGE: ISSUES AND METHODS

NPRDC Special Report 75-9 April 1975

LONGITUDINAL STUDIES OF ATTITUDE CHANGE: ISSUES AND METHODS

Prepared by

Laurie Broedling
Martin Wiskoff
James Herbert

Reviewed by
E. P. Somer

Approved by
James J. Regan
Technical Director

Navy Personnel Research and Development Center
San Diego, California 92152

FOREWORD

This paper represents completion of efforts undertaken to provide a state-of-the-art review of longitudinal attitude assessment methodologies for Technical Panel UTP-1 (Social Processes and Values) in Subgroup U (Behavioral Sciences) of The Technical Cooperation Program (TTCP).

It was first prepared for discussion at a joint UTP-1 and UTP-4 (Manning the Military Forces) meeting in January 1974, in London, England. TTCP is an international organization dedicated to achieving the optimum employment of resources for research and development among participating countries.

This paper has since been circulated, commented upon by the members of Technical Panel UTP-1, and revised accordingly.

Contained within the report is a description of the theoretical issues associated with longitudinal research, the designs and statistical methods available, examples of such research, and the practical problems often encountered. It can serve both as an introduction to this field for social scientists, as well as being a useful reference for social scientists more acquainted with longitudinal techniques. The focus is interdisciplinary, the emphasis is on technological base issues which could be utilized by participating TTCP countries in guiding their R&D programs.

J. J. CLARKIN
Commanding Officer

SUMMARY

Introduction

In order to classify a study as longitudinal, it must meet two criteria. First, observations of one or more variables must be collected on the same group of individuals at two or more points in time. Second, the study must either examine changes in one or more variables over time, or it must examine the relation between different variables at different points in time. Longitudinal designs are superior to cross-sectional designs in studies of attitude change because the former make possible the identification of historical or maturational factors, allow for a clearer specification of causal relationships between attitudes and behavior, eliminate between-group error variance, require a smaller number of subjects, and are more applicable to studies where experimental intervention is required. Longitudinal designs have been usefully applied to analyze many questions of attitude change in military settings.

Theoretical Issues Associated with Studies over Time

The theoretical issues associated with longitudinal designs all cluster around one central issue, that of separating true change from measurement error. Specific issues to be contended with include the over-correction/under-correction problem, the unreliability-invalidity issue, the physicalism-subjectivism dilemma, the contribution of response uncertainty to the total variance, the clearcut conceptualization of what actually constitutes true change in a given study, and the psychometric qualities of the measuring instruments.

Longitudinal Research Designs and Statistical Techniques for Assessment of Attitude Change

Strict experimental designs for accomplishing longitudinal research include the pretest-posttest control group design, Solomon Four-Group design, and a longitudinal sequences design. Appropriate quasi-experimental designs include time-series and panel techniques. Statistical techniques for assessment of attitude change include computing the reliability of a change score, path-analysis, methods based on multiple regression, vector analysis, cross-lagged correlational analyses, dynamic correlational analyses, path-diagram models, and methods based on Markov processes.

Practical Considerations in Doing Longitudinal Research

There are a number of practical problems commonly associated with applying longitudinal techniques. These problems include difficulties in tracking subjects, subject attrition, subject reactivity, the Hawthorne effect, and response to "demand characteristics" of the study.

Conclusions

Advances in longitudinal methodology have been and continue to be made to the point where it is a sophisticated and complex technology. These advances have come from a number of different disciplines. Longitudinal methodology would be applicable in many instances of military studies on attitude change.

TABLE OF CONTENTS

LIST OF FIGURES

LONGITUDINAL STUDIES OF ATTITUDE CHANGE: ISSUES AND METHODS

INTRODUCTION

It has been stated (Parnes, 1972) that in order to classify a study as longitudinal, it must meet the two following criteria: First, observations of one or more variables must be collected on the same group of individuals at two or more points in time. This condition is common to all longitudinal research. Second, the study must either examine changes in one or more of these variables over time, or it must examine the relation between different variables at different points in time. An example of the former type of study is one which measures changes in attitudes toward specific aspects of military life and career intentions from date of entry to the end of recruit training (Schneider & Katz, 1973). The latter is exemplified by research which measures relationships between variables such as interpersonal values, obtained 5-10 weeks after date of service entry, and career motivation, assessed 15-20 weeks after date of entry, by asking the respondents their career intentions (Gordon & Medland, 1964).

Another dimension, and one that applies to both types of studies, is whether the method of investigation is primarily active or passive, i.e., is experimental intervention employed or is the design descriptive? Experimental intervention involves the actual introduction of treatments and measurement of their effects, while descriptive studies are used to assess attitude change as it occurs in the "natural" setting by showing relationships through correlational analysis. Examples of intervention studies are those of Hand and Slocum (1972) and Carron (1964), which were designed to assess managerial improvement programs in industry.

A typical descriptive investigation in the military setting is one which determines "critical periods" of attitude change on variables such as career motivation (Schneider & Katz, 1973; Katz & Goldsamt, 1971; Nelson & Berry, 1966).

Longitudinal versus Cross-sectional Studies of Attitude Change

Some investigators (Horn & Cattell, 1966; Schaie, 1958; Schaie, Rosenthal, & Perlman, 1953) have employed the cross-sectional

approach to measuring change. This technique first involves selecting a stratified random sample of the population consisting of subsamples (cohorts), each representing a level or value of exposure to some independent variable (e.g., time in boot camp for recruits). Then a dependent variable (e.g., attitudes toward the military) is measured for each cohort at the same point in time. The difference scores between cohorts are supposed to reveal a process of attitude change which would be expected to occur in any single group of individuals as they normally progress through the levels of exposure to the independent variable (e.g., changing attitudes that would be expected to occur in a single group of recruits progressing in time through boot camp).

It has been argued (Baltes & Nesselroade, 1972) that the results derived from cross-sectional methods yield incomplete indicators of the occurrence of true change. The contention is that the difference scores obtained from a cross-sectional analysis fail to separate the relative contributions of variables which are presently influencing attitudes from historical variables.

Conversely a longitudinal method can be used to directly assess and identify the variables influencing attitudes over time, with any historical differences related to the variables under investigation minimized. This kind of longitudinal data contributes to the previously mentioned "critical periods" information. A feature of longitudinal as opposed to cross-sectional methods is that in the former a variable can be introduced at a known point in time and systematically manipulated. Still another reason for preferring a longitudinal design over a cross-sectional is that it can document whether the process of cognitive dissonance is operating. Cognitive dissonance theory (Festinger, 1957) maintains that attitude change occurs in order to conform to or become more consistent with an antecedent behavioral event which is related to the attitude in some way. In other words, cognitive dissonance theory suggests that behavior is more often the cause of attitudes rather than the reverse.

There are also two statistical advantages of longitudinal versus cross-sectional studies which deserve mention. First, because there are repeated measures on the same individuals over time, there is no between-group variance. Second, a smaller number of subjects is required than for a completely randomized cross-sectional study.

Longitudinal studies are not without their difficulties, however, despite their theoretical and statistical superiority to cross-sectional

studies. Practical problems associated with their use are discussed in the Practical Considerations in Doing Longitudinal Research section of this paper.

Importance of Longitudinal Attitude Change Studies in the Military Setting

Applebaum and Blakelock (1969) have given three major benefits of longitudinal attitude change studies in the military setting. First, management can be supplied with information showing specific attitude factors which influence changes in other attitudes. For instance, a negative attitude toward supervision might engender a negative attitude toward the military in general. This type of information is valuable for management to consider in making policy decisions. Second, management can view the crystallization of attitudes toward conditions of military life, i.e., from being poorly differentiated to focused on specific aspects of service. Third, longitudinal attitude change research makes it possible to better determine the points in time when career decisions are made, and the relative contribution of attitudinal factors in that decision. Curtis (1973) has pointed out that previous studies on the relationship between attitudes and career decisions have typically measured only postdecision attitudes. He suggests that this practice should be re-evaluated in view of evidence supporting cognitive dissonance theory because postdecision attitudes might be a result, not a predeterminant of a career decision.

The longitudinal technique may also be profitably used to assess attitudes toward different training techniques. Neidt and Meredith (1966) used a longitudinal method to determine the changes in attitudes of a single group of Air Force personnel in training toward a programmed instruction method. First, the personnel were exposed to a conventional training method, after which their attitudes toward the method were measured. Then the programmed instruction method was introduced, and the resulting attitudes were assessed. Finally, the conventional method was reintroduced and attitudes again measured. It was found that individuals responded more favorably toward the programmed instruction than the conventional technique.

THEORETICAL ISSUES ASSOCIATED WITH STUDIES OVER TIME

A prerequisite for the correct application of longitudinal techniques is an understanding of the theoretical issues associated

with studies over time. There are two faces to this body of theory, one the reciprocal of the other. The first is concerned with identification of change over time; the other, identification of stability over time. While the central issues are common, the approaches do differ somewhat. If a researcher is interested in identifying change in an attribute, then any variance due to measurement error is an unwanted source of variance which must be eliminated or statistically controlled. For example, if a person on Questionnaire 1 says he wants to stay in the Navy, and at the time of Questionnaire 2 he still wants to stay in the Navy, but accidentally marks the wrong response on the questionnaire, his results will be misinterpreted. For proper interpretation, the possibility of random measurement error must be built into the study. If a researcher is interested in identifying stability, such as in a study of test-retest reliability, then any true changes in the attributes over time would be designated as random measurement error, and any true change which was present would be the source of error variance in such a study (e.g., Broedling & Mohr, 1973). Taken as a whole, the fundamental theoretical problem in longitudinal studies is partialing out true change from measurement error.

Bereiter (1963) has discussed three issues which have been seen for many years as dilemmas in the measurement of change. The first, the over-correction/under-correction problem, represents the fact any correlation between initial scores and gain scores will be spuriously low. This artificial negative component of the correlation is mathematically due to the fact that the same measurement errors enter into a multiplicative relationship but with opposite sign, resulting in a negative error variance term, as follows:

$$C_{x(y-x)} = C_{(x_t + e_x)(G_t + e_y - e_x)} = C_{x_t G_t} - S^2_{e_x}$$

where $C_{x(y-x)}$ = covariance between initial score X and gain score Y - X,

G_t = true gain score,

e = measurement error,

$S^2_{e_x}$ = error variance of initial scores.

To express this problem another way, if one measures something and that observation is lower than the true score, the gain score measured later will be spuriously higher, that is, the observed gain score will be higher than the true gain score, and vice versa. Consequently, the correlation or covariance between the observed initial scores and observed gain scores will be spuriously low. Bereiter discussed and evaluated a number of methods for controlling this problem, the primary component of these methods being the application of an over-correction which balances the under-correction. Bereiter concluded that there is no complete resolution to the dilemma but that the methods discussed make it possible to reduce the problem to one of choice between reliability estimates.

The second dilemma, the unreliability-invalidity issue, stems from the fact that the higher the correlation between the two administrations, the less will be the reliability of the change scores. Bereiter argued that this is in fact no dilemma at all because the important element is that the reliability of the change scores be as high as possible, even if the correlation between the two administrations decreases.

The third dilemma, physicalism-subjectivism, arises because change, as it is indicated by scale scores, does not necessarily occur in equal intervals. In other words, if between Time A and B, a change of 5 units downward was measured, and between Time B and C, 5 more units downward, it does not necessarily follow that there was an equal amount of change downward in that attitude during the two time periods because the attitude itself may have different unit sizes at different points in the scale. While this dilemma is not resolvable according to Bereiter, he advocated the use of subjective scales over objective scales to help mitigate the problem. Subjective scales are based on people's perceptions of the attributes being measured, while objective scales are based on physicalism, that is, on the objective physical properties of the attributes. Subjective scales produce more interpretable results, particularly when individuals in a study have different initial standings on the attribute being measured.

Coleman (1964) has made an important conceptual contribution to the partialing of change which occurs over time into that which is due to true change and that which is due to measurement error. He maintained that in studies over time, there are three components which contribute to the total variance: true change, measurement error, and response uncertainty. Response uncertainty arises from

the fact that at any given moment a person might be wavering between two or more opinions and might give any one of those opinions as his response to a question on that topic. This component has not ordinarily been made a part of models for longitudinal research, yet it is obviously an important variable in the area of attitudinal research.

An issue which is central to the theory underlying the study of attitude change over time is in the conception of what constitutes change. And what constitutes change is dependent upon how one conceives the characteristic being studied. One well-known distinction is that of trait versus state. The model which would be used to measure changes in a trait, which is a stable characteristic, would be different from the model to measure changes in a state, which is expected to change over time. Cattell (1966) has discussed in length a more detailed breakdown of ways of characterizing change phenomena, and he has developed the following five categories: (1) change conceived as occurring in otherwise "fixed" traits, (2) change in levels of temporary states, (3) change in the environment or environmental relations of the person or group, (4) change evaluated as "tendency to change," i.e., as stability or instability, and (5) change as a characteristic configurational sequence or process.

There is also the possibility that the psychometric qualities of a measuring instrument may change over time. Instrument decay can occur, that is, actual structural changes may take place in the measurement capabilities of the instrument itself. It is not a common problem associated with the questionnaire as a measuring device. Second, changes within the individual which occur between the pretest and posttest, such as learning, can significantly moderate a subject's difference score. Thus, these two factors should be taken into consideration when difference scores are being evaluated.

LONGITUDINAL RESEARCH DESIGNS AND STATISTICAL TECHNIQUES FOR ASSESSMENT OF ATTITUDE CHANGE

This section addresses some recent applications of longitudinal techniques to the attitude change realm, and a brief examination of some statistical methods for the estimate of "true" change. The discussion is divided into methods which can be classified as experimental research designs and those which

represent statistical techniques for assessment of attitude change. The research designs are presented in terms of Campbell and Stanley's (1963) classifications of experimental and quasi-experimental designs. The "change" statistics that will be mentioned are the reliability of difference scores (McNemar, 1969), multiple regression procedures (Cronbach & Furby, 1970), path analysis (Heise, 1969), Markov methods (Coleman, 1964), vector analysis (Carron, 1964), and cross-lagged and dynamic correlations (in Schneider & Katz, 1973; Lawler, 1968; Goodman, 1973).

Experimental Design

To date, relatively few longitudinal attitude change studies have employed true experimental intervention. This is probably because of the difficulty in gaining some degree of control over a field situation--a prerequisite for true experimentation. However, those that have been attempted usually introduce an attitude change program, such as human relations training, and measure the consequent change (Carron, 1964). Other studies have attempted to modify attitudes and demonstrate a relationship between the resulting changes and changes in other behavioral criteria, such as managerial effectiveness (Hand & Slocum, 1972). In both studies the same basic design was used--a pretest-posttest control group design (Campbell & Stanley, 1963). The following diagram illustrates the design:

Experimental Group	$O_1 \ldots\ldots x \ldots\ldots O_2$
Control Group	$O_3 \ldots\ldots\ldots\ldots O_4$

where x = the experimental treatment, the independent variable, and O = observations of attitudes, the dependent variable.

In this design, the experimental group (E) receives a treatment and a control group (C), which has been matched to E through randomized selection on all variables thought to be relevant to the study, receives no experimental treatment. The measures O_1 and O_3 are pretreatment measures of, in this case, attitudes (dependent variable), O_2 and O_4 are posttreatment measures of attitudes. Using the matching feature, all relevant differences between the E and C groups are presumably controlled for. However, extreme care must be taken to make all experimental conditions for both the E and C groups as equivalent as possible, or any measured change between them cannot be safely attributed to the treatment. While the pretest-posttest control group design is very simple, it is perhaps the

most practical for use in a field setting where control of extraneous variables is most difficult.

Another experimental design which could potentially be applied in longitudinal investigations of attitude change is the Solomon Four-Group Design (in Campbell & Stanley, 1963). A diagram of this design is as follows:

Experimental Group	O_1.....x...O_2
Control Group	O_3.....O_4
Experimental Group	.....x...O_5
Control Group	.....O_6

where x = the experimental treatment, the independent variable, and O = the observations of attitudes, the dependent variable.

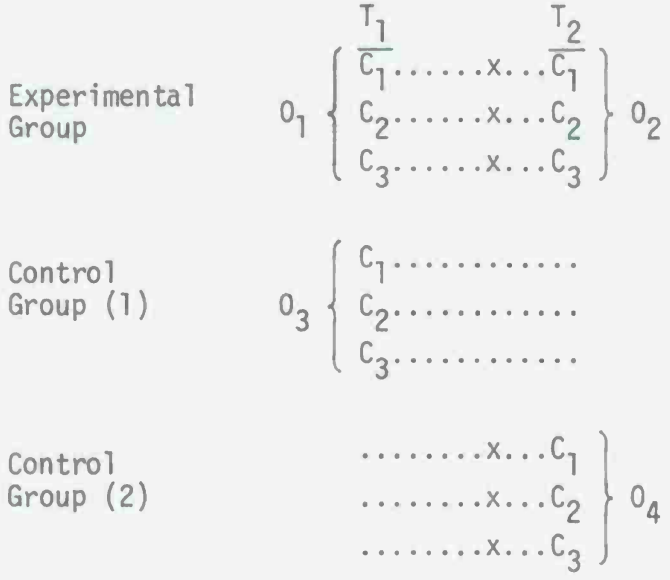
It can be seen from this diagram that the first two levels of this design are equivalent to the pretest-posttest control group design. However, the addition of another experimental and control group, without the pretest, gives considerably more inductive power to the first design. This is because the main effects of testing are separated from the interaction of testing and the experimental treatment (x). For instance, if prejudice was the dependent variable being measured on O_1 by a pretest, and the experimental treatment (x) was a movie on minority civil rights movements, the interaction of O_1 and x may influence the O_2 or posttest measure, and could perhaps be the sole source of any measured change. Thus, the effects of treatment are reproduced in the following four ways:

$O_2 > O_1$	$O_5 > O_6$
$O_2 > O_4$	$O_5 > O_3$

If these relationships are in agreement, inference is increased. In addition, the combined effects of historical and maturational variables can be assessed by comparing O_6 with O_1 and O_3. The decision to use the Solomon Four-Group Design is, of course, a function

of the specific problem under study and resources available to the investigator.

The final experimental design to be reviewed is a unique combination of longitudinal and cross-sectional methods termed "longitudinal sequences" (Baltes & Nesselroade, 1972). Diagrammatically, this design is as follows:



where T = time of testing, O = observation of attitudes (the dependent variable), x = the experimental treatment (independent variable), C = cohorts of the cross-sectional stratification of the groups.

In the above diagram, the experimental group condition consists of a one group pretest-posttest design (Campbell & Stanley, 1963). However, in this case the experimental group has the added feature of being a cross-sectional sample. Control Group (1) is used to assess the potential selective dropout effects, and Control Group (2) to measure any effects of repeated testing. The use of this design makes it possible for the investigator to study problems which require that he separate the changes in attitudes that occur within individuals over time from changes in attitudes that are a result of historical factors unique to each cohort.

Quasi-experimental Designs

In those situations where it is infeasible to maintain the rigor and control of true experimentation, the quasi-experimental designs offer an acceptable alternative. This situation normally pertains to "natural" setting longitudinal research where it is difficult to control treatments applied to groups.

One quasi-experimental method which has been applied in military longitudinal studies of attitude change is the time-series design (e.g., Katz & Goldsamt, 1971). This method involves measuring a single group of subjects at different points in time, on a time continuum, to determine attitudinal trends as a function of the specific associated experience. A general diagram is as follows:

$$O_1 \ldots O_2 \ldots O_3 \ldots O_4 \ldots x \ldots O_5 \ldots O_6 \ldots O_7 \ldots O_8 \ldots$$

It should be pointed out that conclusions based on time-series data are dependent on the magnitude and direction of change occurring between O_1 and O_8. This point is illustrated in Figure 1. If, for example, measures O_1 through O_8 are linear, the gradual change could not be attributed to the experimental variable (x) (A in Figure 1). Conversely, if there is relatively little or no change on measures O_1 through O_4, change between O_4 and O_5 where the treatment occurred, and little or no change between O_5 and O_8, it can more readily be concluded that the obtained change was a result of treatment (B in Figure 1).

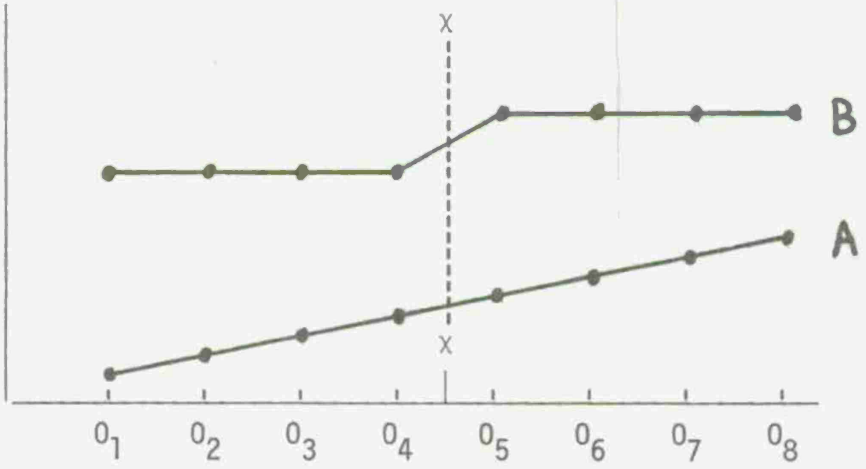


Figure 1. Time-series outcomes.

Another quasi-experimental method used in longitudinal attitude change studies is the panel technique. This approach addresses the question of whether there has been a change in the relationship among certain attitudes, rather than looking at a change in a single attitude. This method entails measuring one group at a minimum of two points in time, perhaps once before the occurrence of an event which is expected to change attitudes and once after that event. At each measurement point, correlations are computed among variables of interest.

Schneider and Katz (1973) used this method to measure changes as a result of recruit training. Before recruit training, measures were made on the recruits' perceptions of the Navy and their intentions to reenlist; these attitudes were intercorrelated to ascertain the relationship between perceptions and intentions. The same steps were taken after recruit training, and then the "before" correlations of perceptions and intentions were compared to the "after" correlations to determine what difference recruit training made in the relationship.

Statistical methods for analyzing panel data are described in the next section. There has also been a recent discussion of the causal logic of panel studies (Howard & Krause, 1970) toward broadening the conceptual framework of analysis of panel data.

Statistical Techniques for Assessment of Attitude Change

Since, as explained in the Theoretical Issues Associated with Studies Over Time section, the assessment of "true" attitude change is contingent upon partialing out the actual change from that which is measurement error, most statistical techniques for measuring attitude change center around this issue.

McNemar (1969) presented a method for computing the reliability of a difference (change) score. The formula is as follows:

$$r_{dd} = \frac{r_{xx} - r_{xy}}{1 - r_{xy}}$$

where r_{dd} = reliability of the difference score,

r_{xx} = reliability of the "after" measure,

r_{xy} = correlation of the "after" measure with the "before" measure.

Once the reliability is known, a standard error of estimate for the change score can be computed and a confidence interval around the score can be constructed.

Heise (1969) developed a technique, based on path analysis, which makes it possible to sort out the contribution of true change and measurement variance to the total variance in a before-after design. This method differs from the traditional test-retest design in that it requires three independent measures of the attitude over time instead of two. The techniques of path analysis entail using the intercorrelations among the various measures to compute the reliability and partial out the variance due to true change. Use of the method requires several assumptions, the validity of which can be tested by taking a fourth independent measure.

Cronbach and Furby (1970) have argued that straightforward change scores, no matter how well adjusted or refined, are rarely useful due to the problems resulting from unreliability of measurement. They also posit that estimation of actual true change is unnecessary to fulfill the purposes of most research. Where it is necessary, they have suggested an entirely different approach for estimating true change using a multiple regression approach. Their equation for the true difference score is:

$$\hat{D}_{\infty} = \beta_1 x + \beta_2 y + \beta_3 w + \beta_4 z + \text{constant},$$

where $\hat{D}_{\infty}$ = estimated true change score,

x = "before" measure,

y = "after" measure,

w and z = demographic variables.

The argument for the addition of weighting by the effect of demographic variables is that it allows for the possibility that any difference in the regression surfaces due to demographic attributes (e.g., males versus females) is taken into consideration. In other words, data on each demographic subgroup are regressed in their own category rather than in the total group data.

As stated earlier, in analyzing attitudes, there is a problem of identifying or untangling attitude changes and response

uncertainties. A model using the concept of Markov processes was developed by Coleman (1964) for the purpose of separating change from the unreliability of response that is sometimes misinterpreted as change. In his book, Coleman illustrates his model with computational methods and a variety of attitude and consumer behavior data; he also provides computer programs written in Fortran to facilitate the computations.

Carron (1964) has developed a method using vector analysis, which takes into account direction as well as magnitude of attitude changes. Vectors on a graph are used to show change from pretest to posttest on some attitude scale. The length of the vector is proportional to magnitude, while its angle with respect to a coordinate system indicates direction.

Schneider and Katz (1973) used a cross-lagged correlational analysis to measure change in panel study data. This technique involves gathering data on two variables that are being tested for causal relationship (e.g., perceptions and intentions), at two points in time (T_1, T_2). From this data, six correlation coefficients are computed (see Figure 2). Correlations (3) and (4) are test-retest, or stability, coefficients for perceptions and intentions, while (1) and (2) are static correlations and provide an indication of covariation among the variables. Correlations (5) and (6) provide the evidence for a causal relationship if:

$$5 > 6 \text{ and } 5 > (1 = 2) > 6$$

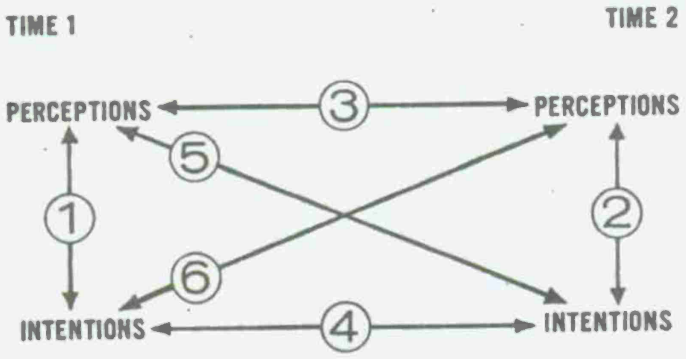


Figure 2. Basic cross-lagged correlation model.

This pattern of correlations would demonstrate that the effects of perceptions on intentions is greater than the effects of intentions on perceptions, and a time lag exists in this causal relationship such that the effects of perceptions on intentions is not immediate. However, if the results show that:

$$6 > (1 = 2) > 5,$$

then there is reason to believe that intentions caused perceptions. In order to rule out the possibility that a third (not measured) variable was not responsible for the obtained causal relationship, a dynamic correlation (Lawler, 1968) is computed for both perceptions and intentions. This involves correlating the difference between perceptions at T_1, and T_2 (ΔA), and the difference between intentions at T_1 and T_2 (ΔB). Therefore, changes in one variable are being correlated with changes in the other. In order for a dynamic correlation to be false, a third or unmeasured variable must vary in different amounts or directions in the sample members, and these variations would have to be highly correlated with both ΔA and ΔB. It is unlikely that any variable would meet this criterion, so most third causal variables can be ruled out.

Goodman (1973) has also recently developed new models and methods for the causal analysis of data obtained in panel studies through the use of path-diagram models. First, he provides a method for estimating the magnitudes of various effects between variables, represented by path diagrams. Second, he presents methods such as the χ^2, or goodness-of-fit statistic, for determining whether a given path-diagram model (as described by a given set of equations) is an adequate representation of the observed data. Third, an overall test of the whole system of equations in the path-diagram model, as well as tests for each separate equation, are provided. Fourth and last, methods for comparing the advantages and disadvantages of a variety of path-diagrams describing the data are presented.

PRACTICAL CONSIDERATIONS IN DOING LONGITUDINAL RESEARCH

In addition to the theoretical and statistical aspects of employing longitudinal designs, there are practical concerns as well. One requirement is that subjects' whereabouts be documented to maximize the percentage of subjects available for participation at all administration points. If subjects are lost from the study simply because they cannot be located, it may have damaging effects on the accuracy of results. It is possible that such losses

introduce a specific form of nonrespondent bias, that is, people who cannot be located differ in some systematic way from those who can be located. Such a situation may often exist in military settings where certain subgroups of people tend to move less often than others and where better administrative records are kept on some than on others (e.g., high-ranking officers versus low-graded enlisted). Actual experience with trying to track naval personnel in longitudinal studies has shown that use of existing data files, such as personnel master tapes, is not adequate (Goldsamt, 1973). It is far better to establish a tracking system specific to the purposes of each individual investigation.

The typical reaction to subject attrition is to drop all those cases for whom the data are incomplete. As mentioned above, such losses may produce nonrespondent bias, but at the very least, it entails the discarding of data on those subjects who completed some but not all of the questionnaires. In order to make better use of the available data, Lehnen and Koch (1974) developed a model in which all the data can be utilized, including the data from those subjects who were available for only part of the questionnaire administrations.

A second practical requirement is that subjects who participate must not have their responses affected by the fact that they are participating. This problem can take several specific forms, such as the Hawthorne effect, reactivity in social research, and results of the use of obtrusive measures. This problem is by no means unique to longitudinal research, but it is accentuated within longitudinal designs due to the fact that there is more time available for the subjects to become reactive and due to the fact that in longitudinal research measures are repeatedly taken. The Hawthorne effect, of course, is defined as the increase of positive attitudes due to the subjects' awareness of experimental participation. However, in some instances and for certain types of people, the opposite effect occurs. Reactivity in response to repeated measures occurs when a pretest either increases or decreases a subject's awareness of the experimental variable (Campbell & Stanley, 1963). If a pretest does alert the subject as to how he is expected to change his attitudes, this reactivity will confound the results obtained on subsequent testing or measurement. Rosnow and Suls (1970) have hypothesized that reactivity from pretesting may come less from the introductory summary contained in the pretest than from the cues given regarding the manipulatory character of the research. These are known as the demand characteristics of the study, and the results will then be

affected by the willingness of subjects to comply to these "demands." Another facet of this problem is the inadvertent use of a biased questionnaire which actually induces changes in the later attitudes of the subjects (Dillehay & Jernigan, 1970). Concern over reactivity of subjects to psychological research has been growing in recent years, and ideas for coping with the problem have been generated (e.g., Webb, Campbell, Schwartz, & Sechrest, 1966).

However, despite the aforementioned problems, the longitudinal method is viewed as the most promising for measurement of change. With the increasing improvement in unobtrusive measures (e.g., Webb et al., 1966), and with better controls for subject mortality, the longitudinal approach can be an effective research tool.

CONCLUSIONS

Advances in longitudinal methodology have been and continue to be made to the point where it is a very sophisticated and complex technology. It is encouraging to note that the methods are based on theoretical foundations, as well as on statistical and mathematical tools. Unfortunately, this technology has been developed in many diverse disciplines, and much of this development has gone on independently. Also, some theoretical issues pertaining to the use of longitudinal techniques in psychological research remain unsolved. This paper has drawn together longitudinal techniques from several disciplines so that their common methods and problems can be ascertained.

REFERENCES

Applebaum, L., & Blakelock, E. Personnel reactions to incentives, naval conditions and experiences (PRINCE): Study Plan. Washington, D. C.: Naval Personnel Research and Development Laboratory, 1969. (WSS 70-1)

Baltes, P. B., & Nesselroade, J. R. Cultural change and adolescent personality development: An application of longitudinal sequences. Developmental Psychology, 1972, 7, 244-256.

Bereiter, C. Some persisting dilemmas in the measurement of change. In W. Harris (Ed.), Problems in measuring change. Madison, Wisconsin: University of Wisconsin Press, 1963.

Broedling, L. A., & Mohr, E. S. Sources of error in naval personnel surveys: An investigation. Washington, D. C.: Naval Personnel Research and Development Laboratory, 1973. (WTR 73-46)

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Carron, J. Human relations training and attitude change: A vector analysis. Personnel Psychology, 1964, 17, 403-422.

Cattell, R. B. Handbook of multivariate experimental psychology. Chicago: Rand McNally, 1966.

Coleman, J. Response uncertainty and the measurement of change. Englewood Cliffs, N.J.: Prentice-Hall, 1964.

Cronbach, L. J., & Furby, L. How should we measure change?--or should we? Psychological Bulletin, 1970, 74, 68-80.

Curtis, E. Relationship of attitudes and behavior. Unpublished manuscript. Navy Personnel Research and Development Center, San Diego, 1973.

Dillehay, R. C., & Jernigan, L. R. The biased questionnaire as an instrument of opinion change. Journal of Personality and Social Psychology, 1970, 15, 144-150.

Festinger, L. A theory of cognitive dissonance. New York: Harper & Row, 1957.

Goldsamt, M. R. Attitudinal changes in Category IV perceptions of the Navy during recruit training. Washington, D. C.: Naval Personnel Research and Development Laboratory, 1973. (WTR 73-17)

Goodman, L. A. Causal analysis of data from panel studies and other kinds of surveys. American Journal of Sociology, 1973, 78, 1135-1191.

Gordon, L. U., & Medland, F. F. Values associated with military career motivation. U. S. Army Personnel Research Office, 1964. (Technical Research Note 143)

Hand, H. H., & Slocum, J. W., Jr. A longitudinal study of the effects of a human relations training program on managerial effectiveness. Journal of Applied Psychology, 1972, 56, 412-417.

Heise, D. R. Separating reliability and stability in test-retest correlation. American Sociological Review, 1969, 34, 93-101.

Horn, J. L., & Cattell, R. B. Age differences in primary mental ability factors. Journal of Gerontology, 1966, 21, 210-220.

Howard, K. I., & Krause, M. S. Some comments on "Techniques for estimating the source and direction of influence in panel data." Psychological Bulletin, 1970, 74, 219-224.

Katz, A., & Goldsamt, M. R. Assessment of attitudes and motivations of Category IV marginal personnel: Demographic characteristics, attitudes and personal adjustments during recruit training. Washington, D.C.: Naval Personnel Research and Development Laboratory, 1971. (WRR 72-1)

Lawler, E. E. A correlational-causal analysis of the relationship between expectancy attitudes and job performance. Journal of Applied Psychology, 1968, 52, 462-468.

Lehnen, R. G., & Koch, G. G. Analyzing panel data with uncontrolled attrition. The Public Opinion Quarterly, 1974, 38, 40-56.

McNemar, Q. Psychological statistics. New York: Wiley, 1969.

Neidt, C. O., & Meredith, T. F. Changes in attitudes of learners when programmed instruction is interpolated between two conventional instruction experiences. Journal of Applied Psychology, 1966, 50, 130-137.

Nelson, P. D., & Berry, N. H. Attitudes of Marines during first enlistment. San Diego: U.S. Navy Neuropsychiatric Research Unit, 1966. (Report Number 66-21)

Parnes, H. Longitudinal surveys: Prospects and problems. Monthly Labor Review, 1972, 95, 11-15.

Rosnow, R. L., & Suls, J. M. Reactive effects of pretesting in attitude research. Journal of Personality and Social Psychology, 1970, 15, 338-343.

Schaie, K. W. Rigidity-flexibility and intelligence: A cross-sectional study of the adult life span from 20 to 70 years. Psychological Monographs, 1958, 72 (Whole No. 462).

Schaie, K. W., Rosenthal, F., & Perlman, R. M. Differential deterioration of functionally "pure" mental abilities. Journal of Gerontology, 1953, 8, 191-196.

Schneider, J., & Katz, A. PRINCE Report No. 3: Measurement of attitude change during recruit training. Washington, D.C.: Naval Personnel Research and Development Laboratory, 1973. (WTR 73-11)

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. Unobtrusive measures: Nonreactive research in the social sciences. Chicago: Rand McNally, 1966.

DISTRIBUTION LIST

Director of Defense Research and Engineering (Acquisition Management)
Chief of Naval Operations (OP-987E)
Chief of Naval Personnel (Pers-10c)
Chief of Naval Education & Training (N-33)
Chief of Naval Technical Training
Chief of Naval Material (NMAT-030B)
Chief of Naval Education & Training Support
Chief of Naval Research
  (Code 450)
  (Code 458)
Commandant of the Marine Corps (Code MP)
Commander, Navy Recruiting Command
Commanding Officer, Naval Training Equipment Center
Commanding Officer, Navy Medical Research Center
Commanding Officer, Naval Aerospace Medical Institute
Commanding Officer, Naval Submarine Medical Center
Commanding Officer, Naval Medical Research Institute
Commanding Officer, Naval Development & Training Center
  (Code 0120)
Superintendent, U. S. Naval Academy
Superintendent, Naval Postgraduate School
Superintendent, U. S. Military Academy
Superintendent, U. S. Air Force Academy
Superintendent, U. S. Coast Guard Academy
Director, Navy Occupational Task Analysis Program Department,
  Naval Personnel Program Support Activity
Chief of Research and Development, U. S. Army
Army Research Institute for Behavioral and Social Sciences
Personnel Research Division, Air Force Human Resources Laboratory
  (AFSC), Lackland Air Force Base
Occupational and Manpower Research Division, Air Force Human Resources
  Laboratory (AFSC), Lackland Air Force Base
Air Force Human Resources Laboratory, Brooks Air Force Base
Commander, Armed Forces Vocational Testing Group
Defense Race Relations Institute
Center for Naval Analyses

TTCP Distribution

LTCOL H. Taylor, U.S. National Leader, Subgroup U (5)
Mr. E. D. Treharne, Cognizant Washington Deputy
COL A. J. Affleck, Australian National Leader
Dr. C. H. Baker, Canadian National Leader
Mr. R. Speight, Executive Chairman and UK National Leader
Dr. J. E. Uhlaner, U.S. Army Principal Member
Dr. G. Bryan, U.S. Navy Principal Member
Dr. J. McReynolds, U.S. Air Force Principal Member
Dr. D. R. Segal, Chairman, UTP-1
Mr. A. Cassie, UK National Leader, UTP-1
LTCOL M. Martin, Canadian National Leader, UTP-1
LTCOL G. Owens, Australian National Leader, UTP-1
Dr. N. Guinn, U.S. Air Force Member, UTP-1
Mr. J. James, UK Air Force Member, UTP-1
Mr. R. A. Godwin, UK Navy Member, UTP-1

NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER
SAN DIEGO, CALIFORNIA 92152

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
DEPARTMENT OF THE NAVY
DOD-316

U.S.MAIL